Comparison of image registration methods for composing spectral retinal images

Laaksonen, Lauri; Claridge, Ela; Falt, Pauli; Hauta-Kasari, Markku; Uusitalo, Hannu; Lensu, Lasse

DOI:
10.1016/j.bspc.2017.03.003

Comparison of Image Registration Methods for Composing Spectral Retinal Images

Lauri Laaksonen[a], Ela Claridge[b], Pauli Fält[c], Markku Hauta-Kasari[c], Hannu Uusitalo[d,e], Lasse Lensu[a,*]

[a] *Machine Vision and Pattern Recognition Laboratory, School of Engineering Science, Lappeenranta University of Technology, PO Box 20, FI-53851 Lappeenranta, Finland*
[b] *School of Computer Science, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom,*
[c] *School of Computing, University of Eastern Finland, P.O. Box 111, FI-80101, Joensuu, Finland*
[d] *Department of Ophthalmology, University of Tampere, Kalevantie 4, FI-33014 University of Tampere, Tampere, Finland*
[e] *SILK Research and Development Center for Ophthalmic Innovations, School of Medicine, Biokatu 14, FI-33014 University of Tampere, Tampere, Finland*

Abstract

Spectral retinal images have significant potential for improving the early detection and visualization of subtle changes due to eye diseases and many systemic diseases. High resolution in both the spatial and the spectral domain can be achieved by capturing a set of narrow-band channel images from which the spectral images are composed. With imaging techniques where the eye movement between the acquisition of the images is unavoidable, image registration is required. As manual registration of the channel images is laborious and prone to error, a suitable automatic registration method is necessary.

In this paper, the applicability of a set of image registration methods for the composition of spectral retinal images is studied. The registration methods are quantitatively compared using synthetic channel image data of an eye phantom and a semisynthetic set of retinal channel images generated by using known transformations. The experiments show that generalized dual-bootstrap iterative closest point method outperforms the other evaluated methods in registration accuracy, measured in pixel error, and the number of successful registrations.

Keywords: Image registration, Spectral imaging, Retinal imaging, Fundus imaging, Quantitative evaluation

1. Introduction

Eye diseases such as diabetic retinopathy (DR), glaucoma and age-related macular degeneration (AMD), or complications of many systemic diseases like diabetes and systemic hypertension (SH), cause structural changes in the eye fundus. Early detection of the retinal changes, monitoring of their progress and risk factor analysis allow better and more cost-effective treatment as most diseases can be successfully treated if they are diagnosed early and monitored regularly (e.g., [1, 2]).

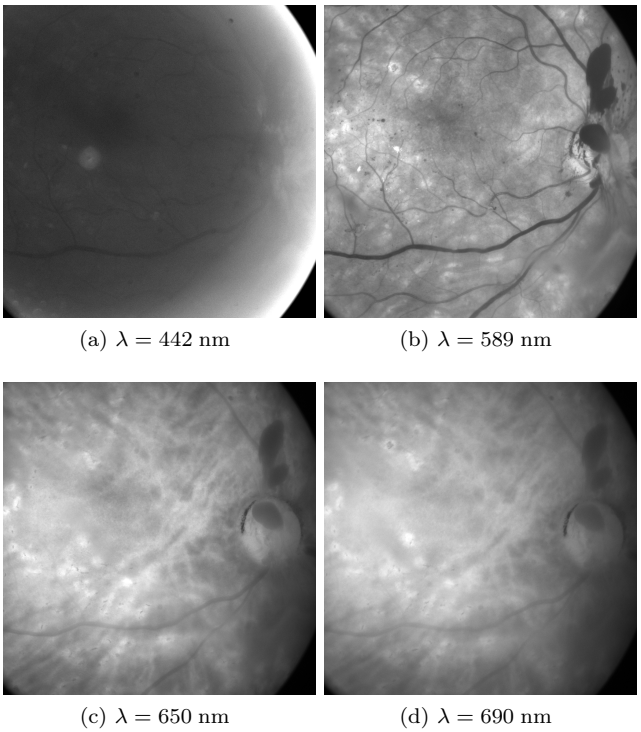
Retinal imaging provides a non-invasive view into the eye and its vascular bed. It is the standard practice to screen, diagnose and monitor eye diseases. Greyscale or RGB images with high spatial resolution are commonly used in the diagnosis, complemented with more advanced eye imaging methods when necessary. To support further development of the diagnostic tools, methods for spectral reflectance measurements, especially spectral imaging of the retina, have been developed [3, 4, 5, 6, 7, 8]. The aim of the development has been to improve the capabilities to detect and visualize different parts of the retina and lesions related to the eye diseases.

Depending on the imaging technology, composing high-resolution spectral images based on a set of channel images may require image registration. In [4], for example, an imaging system for capturing spectral images of the retina using a set of 30 narrow-band interference filters is presented. Changing the filters requires time, thus, the eye moves with respect to the camera and image registration is needed. The purpose of the registration is to find the geometric transformation needed to spatially align the images with each other. Manual registration by selecting corresponding points in image pairs becomes difficult and even infeasible when the number of individual channels increases, or when the image features become less salient between the images. Four example images acquired with the imaging system described in [4] using 2×2 pixel binning are shown in Fig. 1.

To solve the image alignment problem, there exists a significant body of work in the field of image registration. However, the majority of the approaches are designed for images originating from the same imaging modality. While registration approaches designed for multimodal data exist (e.g., [9, 10]), the resulting images are expected to be similar enough for feature matching. Neither of these prerequisites are necessarily true for the channel images

*Corresponding author at: Machine Vision and Pattern Recognition Laboratory, School of Engineering Science, Lappeenranta University of Technology,PO Box 20, FI-53851 Lappeenranta, Finland, Tel.: +358 294 462 111; fax: +358 5 411 7201.

Email address: lasse.lensu@lut.fi (Lasse Lensu)



(a) $\lambda = 442$ nm (b) $\lambda = 589$ nm

(c) $\lambda = 650$ nm (d) $\lambda = 690$ nm

Figure 1: Example channel images captured with the system in [4]. The images are normalised to the intensity mean of 0.5 and standard deviation of 1 for visualisation.

when the difference in acquisition wavelength is large. Despite the fact that in the case of spectral retinal imaging the channel images are captured with a single modality, the image appearance varies significantly as can be seen in Fig. 1.

This paper presents a comparison of image registration approaches for composing spectral retinal images from channel images, and more generally, for aligning retinal images with significantly different visual information content. The compared methods are a well-performing subset of the methods evaluated in the preceding study [11]. As in the previous study, this paper quantitatively evaluates the methods' performance on synthetic and semisynthetic retinal image data. Here the comparison is significantly deepened by including the following subtopics: i) the effect of different image set registration strategies on the spatial accuracy of alignment, ii) the residual spatial inaccuracy and its effect on the spectral image quality and iii) presentation of the evaluation results more inclusively.

2. Related work

Image registration in general and in the field of medical imaging is a widely studied problem. An example review of general image registration methods based on both features and similarity metrics has been presented by Zitova and Flusser [12]. The review includes discussion on approaches to feature detection and matching, mapping function design, and image transformation and resampling. The evaluation of registration performance of feature and area based registration methods is also discussed. More modern approaches to image registration are presented in the study by Wyawahare et al. [13].

A significant part of medical image registration literature focuses on magnetic Resonance Imaging (MRI), computed tomography (CT) and other radiological modalities. Maintz and Viergever [14] present an extensive review of medical image registration approaches. A significant number of the methods reviewed deal with the registration of radiological images, with methods dealing with monomodal, multimodal and modality to model registration. Hill et al. [15] review the main approaches for registering radiological images. The review presents an overview of rigid feature-based methods and intra- and intermodal (voxel) similarity-based methods. Bhatia et al. [16] present a qualitative evaluation of similarity metrics for groupwise non-rigid registration, including a novel metric. The methods are evaluated on MRI data.

A more recent review on medical image registration is the one by Markelj et al. [17]. Three classes of registration base and strategy are identified: feature-, intensity- and gradient-based methods, and projection, back-projection and reconstruction strategies. However, while the paper cites a number of other modalities, the scope of the review is limited to 3D-to-2D registration. In [18], Oliveira and Tavares describe the geometric transformations, similarity measures and optimisation methods in common (medical) registration approaches. In addition, available registration software and methods for performance evaluation are reviewed.

Deformable transformation models have been widely used in medical image registration to cope with dynamically changing organs and inter-person variation in the anatomy. Crum et al. [19] present an overview of deformable medical image registration, with the presented methods dealing mostly with radiological modalities. A number of well-known similarity measures, such as sum of squared differences (SSD), correlation coefficient (CC) and mutual information (MI) are included. Non-rigid transformation models including splines and demons are also discussed. Bhatia et al. [16] present a qualitative evaluation of similarity metrics for group-wise non-rigid registration, including a novel metric. The methods are evaluated on MRI data. Sotiras et al. [20] present a comprehensive study of recent approaches to deformable image registration. A large number of deformable registration methods, classified by the deformation models, matching criteria and optimisation approach used, are described. While not limited to the application area, the study puts an emphasis on methods dealing with the registration of medical images.

Spectral imaging typically produces images with tens or hundreds of spectral channels. Compared to the traditional three-channel red-green-blue (RGB) or single-channel grayscale images, spectral images include a superiorly vast and detailed information content encoding, for example,

accurate spectral color information of an imaged object. Methods for registering spectral images have been proposed by, e.g., [21, 22, 23]. However, the papers on the registration of spectral images focus on remote sensing data which are commonly acquired from a large distance. In retinal imaging, due to the wide-angle optics and the significantly curved imaging target, the pose changes of the eye w.r.t. the camera cause warping in and distortions near the edges of the images. In addition, the penetration depth of light into the multilayered tissue of the eye fundus differs at different wavelengths which causes significant variation in the channel images.

While a large part of medical image registration literature is focused on radiological modalities, methods for registering retinal images have also been studied. As a part of their review on methods applicable to the automatic screening of diabetic retinopathy, Teng et al. [24] present an overview of feature-based registration methods and two methods that utilise the whole retinal image. The features used include matched filter responses, vessel branching points and manually marked anatomy markers. The reviewed methods were constrained to rigid transformation models. Laliberté et al. [25] quantitatively evaluate registration methods on retinal colour and fluorescein angiography images. A novel method based on vessel network structure is also presented.

While a significant body of work related to image registration methods and their application in medical imaging and spectral images, little attention has been given to registering retinal images captured with different narrowband illuminations. Such registration is required, e.g., to compose spectral retinal images from channel images and to align retinal images originating from different modalities for fusing the image information.

3. Methods

The purpose of pair-wise image registration is to find the geometric transformation needed to align a floating image (i.e., the image to which the registration transformation is applied) with the base image (i.e., the registration target). This section shortly describes the registration methods included in the comparison. The methods are a subset of the ones evaluated in [11], and they were selected based on their performance with similar data. Publicly available implementations [26, 27, 28, 29] were used for all the methods.

3.1. Registration by local similarity

In local similarity based registration, the floating image is deformed so that the set of local similarities is maximized. In the framework used here, the nodes of an $n \times n$ grid are iteratively moved based on local similarity of the base and floating image. The final transformation is obtained by interpolation with b-splines using the grid nodes as control points.

The similarity measures MI [9] and minimization of residual complexity (RC) [30] were selected for the comparison. MI [9] is a measure of similarity (or dependence) between two datasets, measured as the distance between their joint probability distribution and the independent probability distribution. Methods based on maximization of mutual information have seen frequent use in registration of multimodal medical images.

RC [30] is a similarity measure that accounts for spatial intensity distortions and is based on the minimization of the complexity of residual image. The method outperforms state-of-the-art similarity measures in several medical registration problems (including retinal image registration), but is limited to monomodal data.

3.2. Feature-based registration

The generalized dual-bootstrap iterative closest point (GDB-ICP) [31] algorithm finds a transformation aligning two images by starting from a bootstrap region (small area of overlap) between the images and a locally stable similarity transformation. An initial transformation derived from a scale-invariant feature transform (SIFT) descriptor match is refined and validated by feeding edge and corner points inside a growing bootstrap region to a robust iterative closest point (ICP) algorithm.

To reduce the number of incorrect feature matches, the edge and corner points are divided into driving and matchable features. The driving features with stricter validation thresholds are matched to a larger pool of matchable features. To increase the bootstrap region stability, GDB-ICP determines both the forward and inverse transformations, and uses bi-directional mapping of the feature points to increase the number of constraints.

4. Registration strategy

In the case where a set of images need to be aligned with each other, different strategies can be taken to produce the complete alignment result. In spectral image composition, all channels need to be transformed into the same space to form a full spectral image. This section discusses different strategies for determining the set of pairwise registrations to compose a full spectral image where all channels are aligned.

4.1. Registration to a single base image and sequential registration

A simple approach for aligning a channel image set is to register each image to a previously selected base image. For spectral images, however, registering all channel images to a single base image means that image features change considerably as a result of large differences in illumination wavelength between the floating and base image. The difference in the illumination wavelength significantly affects their structure and intensity due to varying reflectance of different features of the retina.

Large differences in image structure and intensity due to wavelength can be avoided by registering each channel image to its immediate neighbor along the spectral dimension. For images that are not immediate neighbours of the base image, the transformations steps can be sequentially combined to align the images with the chosen base image.

4.2. Joint registration

Instead of registering the channel image set as a number of independent pair-wise registrations, the overall registration result is likely to be improved if the registration strategy consideres the whole image set. Joint registration has been extensively studied, especially in MRI and tomography [32, 33, 34, 35, 36, 37]. However, the majority of joint registration approaches in the literature are applicable (at least without significant changes) only to methods that make use of a similarity measure, and are often an integrated part of the method. Considering the promising performance of feature-based methods in retinal image registration, two joint registration strategies applicable to both feature and similarity measure based registration are studied.

4.2.1. Registration using intermediate templates

Instead of directly registering to a neighbouring or to the base image, the floating image can be registered to a template to avoid both the accumulation of error due to combined transformations and large differences in features and intensities. Each floating image is registered to a template that is a combination of the previously registered intermediate (i.e., channels between the floating and the base) images. Similar approaches for groupwise registration have been presented in e.g., [38, 39].

The template T_i used as the registration target for floating image I_i can be defined as

$$T_i = \frac{I_{i-1} + T_{i-1}}{2}, \tag{1}$$

where T_{i-1} is the template of the previous step and I_{i-1} is the previously registered floating image. In the first step, the base image is used as the template. Each intermediate image can be given an equal or unequal weight in the generation of the template. As each template has been (not accounting for the registration error) transformed into the same space with the base image, the resulting transformation is of the same complexity for each channel image.

4.2.2. Registration using shortest path

The sequential registration strategy can be improved by registering only the images along the shortest path from a floating image to the base image instead of all the intermediate images. Here, the cost of each registration step is measured in image similarity. A cost matrix is defined by calculating the squared sum of intensity error for each combination of pairs for the image set. For each image, Dijkstra's algorithm is applied to the cost matrix to determine the shortest path to the base image. The final transformation is determined by combining the pair-wise transformations along the shortest path.

5. Experiments

5.1. Datasets and performance evaluation

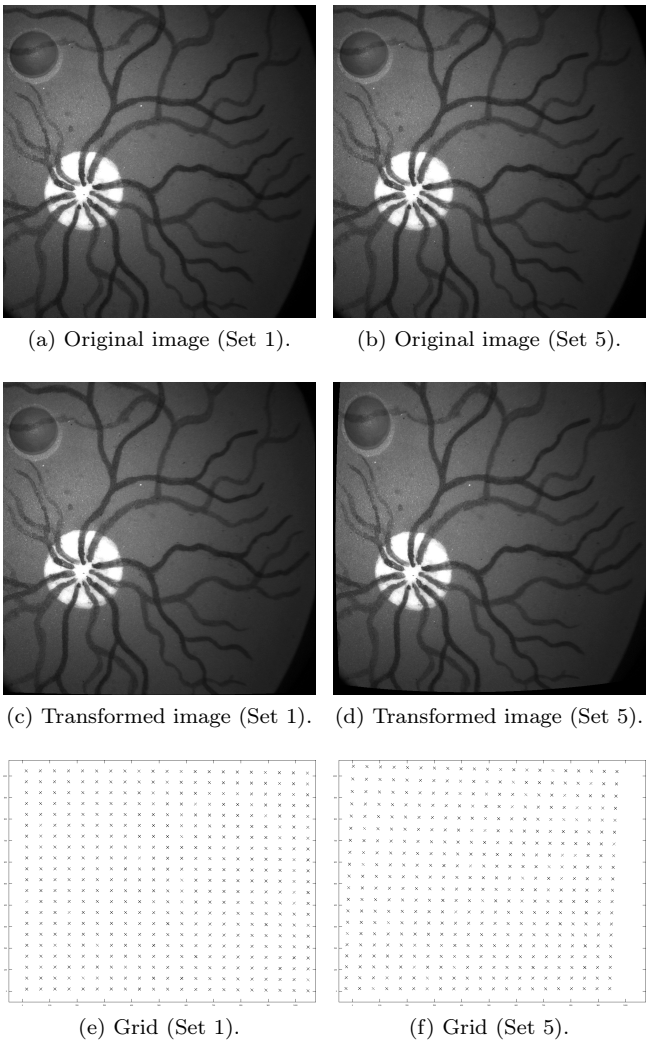
To quantitatively evaluate the performance of each image registration method, two test sets were generated: a synthetic set of spectral images of an eye phantom (artificial eye), and a semisynthetic set of in-vivo images of the human retina where the channel image alignment is limited only by the small system error related to image capture.

The synthetic test set was based on spectral images of a Carl-Zeiss Meditec eye phantom acquired with the system described in [4]. The phantom is a closed container with a small entry pupil fitted with an optical system simulating the lens of the human eye. The back of the container is concave (to represent the curvature of the eye fundus), with painted retinal structures (e.g., vasculature and fovea). The imaging system is composed of a 45-degree Canon CR5-45NM fundus camera, Schott Fostec DCR III fiber optic halogen source, a set of 30 bandpass interference filters spanning the wavelength range of $400 - 700$ nm at approximately 10 nm intervals and average full width at half maximum of 10 ± 2 nm, and a QImaging Retiga-4000RV grayscale digital CCD camera with the sensor resolution of 2048×2048 pixels. A fully aligned spectral image consisting of 30 channels with spatial resolution of 1024×1024 (2×2 pixel binning) of the phantom was used as the basis. Example images acquired with the setup are shown in Fig. 1.

The semisynthetic test set was based on spectral retinal images captured with the system described in [40]. The system consists of a modified Zeiss fundus camera (RCM250) synchronised with a Hamamatsu cooled EM-CCD (ImageEM C1300-13). The retina is illuminated at six selected wavelengths $507, 525, 552, 585, 596, 611$ nm [8] using a halogen white-light source (OSL1, Thorlabs Inc, Newton, NJ, USA) filtered through a liquid crystal tunable filter (VariSpec, CRI, U.S.A.). The average acquisition time is approximately 0.5 seconds per one spectral image [40]. During the experiments, the maximum inter-channel displacement was found to be 2.3 pixels (referred to as the system error).

For both the synthetic and semisynthetic test sets, five image sets were generated from the original images by transforming each channel by a designed transformation. To simulate the changes in retinal images as the eye moves in relation to the camera, the test set images were first projected onto a hemisphere with dimensions corresponding to the average human eye (as reported in [41]). The hemisphere was translated and rotated to simulate an offset in the optical axis and movement of the eye with respect to the camera. The image coordinates were then projected back onto a plane, and the deformed image was gained by

estimating the values at the back-projected coordinates by bicubic interpolation. Examples of the images in the synthetic and semisynthetic test sets, and the corresponding transformations are shown in Figures 2 and 3.



(a) Original image (Set 1). (b) Original image (Set 5).

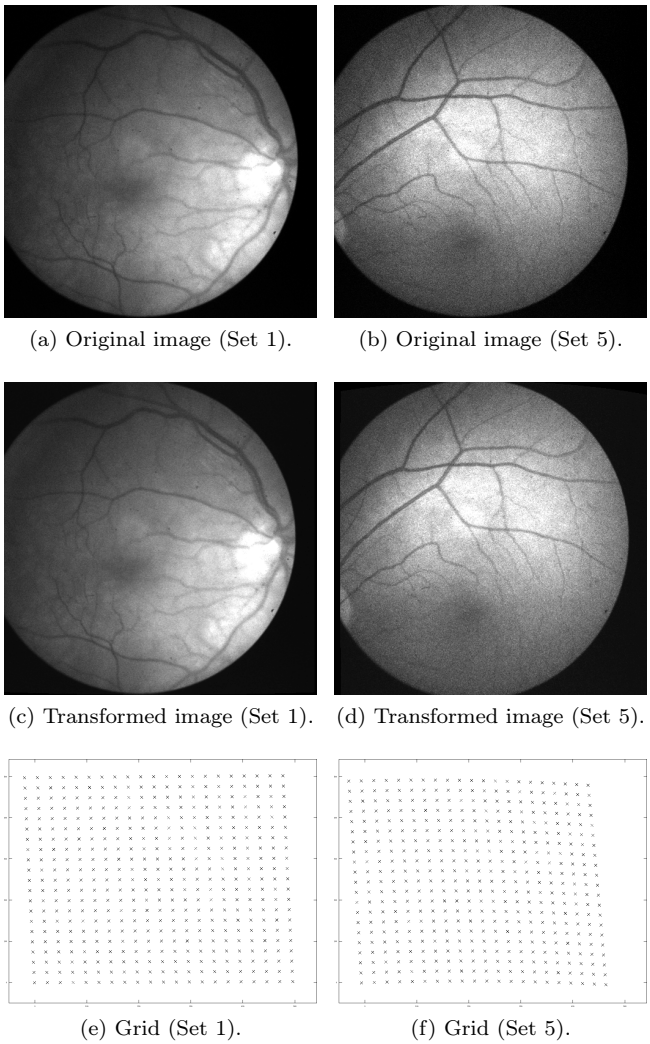
(c) Transformed image (Set 1). (d) Transformed image (Set 5).

(e) Grid (Set 1). (f) Grid (Set 5).

Figure 2: Synthetic test set based on eye phantom images; examples at $\lambda = 589$ nm of the original and transformed images (enhanced for visualization) with corresponding transformation grids.

To validate the approach for generating the test sets, images captured with a second eye phantom setup, for which the angle between the phantom and camera can be accurately set, were deformed with the approach for test set generation. The phantom consists of a planar target, with retinal features painted on the surface, mounted on a frame that can be rotated in small increments in two dimensions. The validation set contained three sets of images with the phantom in 6 different rotations in relation to the camera (3°, 1.5°, −1.5°, −3°, −4.5°, −6°). An example is shown in Fig. 4.

An image with the phantom directly facing the camera was deformed with an increasing angle θ, and the deformed image was compared to an image with the phantom physically rotated θ degrees. The error between the physical and simulated deformation was determined as the displacement of speeded-up robust feature (SURF)-keypoints [42] visible in both images. The numerical results are shown in Table 1.



(a) Original image (Set 1). (b) Original image (Set 5).

(c) Transformed image (Set 1). (d) Transformed image (Set 5).

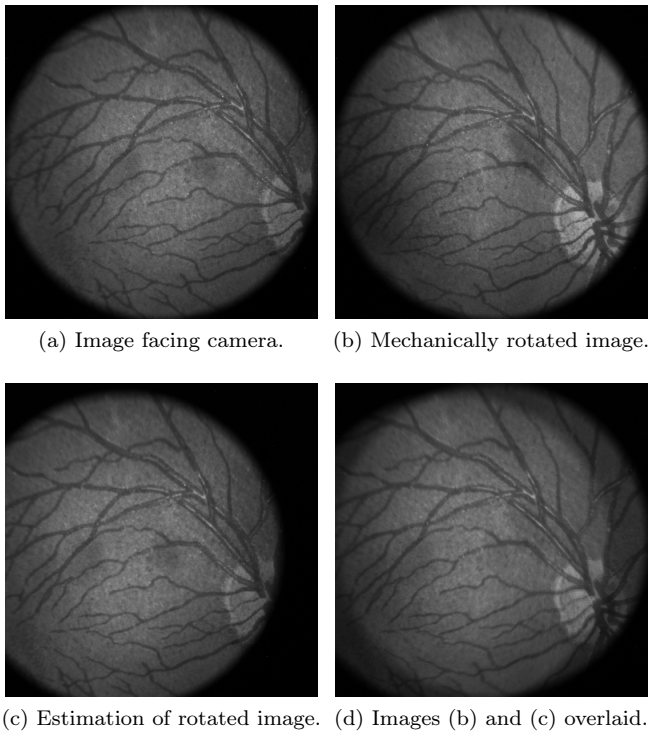
(e) Grid (Set 1). (f) Grid (Set 5).

Figure 3: Semisynthetic test set based on retinal images; examples at $\lambda = 585$ nm of the original and transformed images (enhanced for visualization) with corresponding transformation grids.

Table 1: Validation of the synthetic data generation approach; errors (in pixels) between feature locations in mechanically rotated and artificially deformed images, respectively.

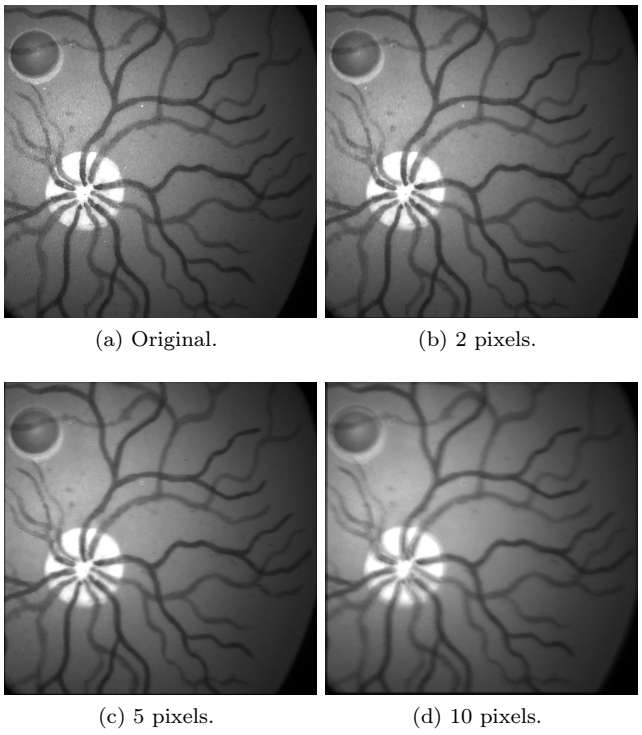
θ	−6°	−4.5°	−3°	−1.5°	1.5°	3°
Mean	2.98	1.85	1.38	0.83	1.71	2.74
Std	0.67	0.48	0.45	0.32	1.00	1.07

It should be noted that the artificial retinal features of the phantom used for validating the synthetic data generation approach are painted on a plane (note: different eye phantoms were used for the synthetic test set and for the validation of the synthetic data generation). Conse-



(a) Image facing camera. (b) Mechanically rotated image. (c) Estimation of rotated image. (d) Images (b) and (c) overlaid.

Figure 4: Examples of the phantom set used for validating synthetic data generation.



(a) Original. (b) 2 pixels. (c) 5 pixels. (d) 10 pixels.

Figure 5: Channel-wise mean images of the spectral quality reference; each channel of the spectral image is translated n pixel to a different direction.

quently, the pixel locations of the deformed validation set images are projected onto a plane instead of a hemisphere.

Due to the differences between the true retinal curvature (semisynthetic data), the curvature of the synthetic test set phantom and an ideal hemisphere, the deformations in the synthetic ground truth are unlikely to correspond as closely with the eye movements they are simulating. However, the purpose of the deformation approach is not to accurately simulate specific eye movements, but to produce realistic deformations to the synthetic ground truth. Based on the validation results, the approach to synthetic ground truth generation can be expected to produce deformations sufficiently similar to those caused by the movements of the eye in relation to the camera.

The channel-wise registration errors can cause significant error in the image spectra when individual pixels are considered. To estimate the detoriation of the quality of the spectra as the registration error increases, artificial systematic misalignment was applied to a spectral image of the eye phantom and metrics measuring the quality of the spectra were calculated.

30 channel images of the eye phantom were translated n pixels in a direction unique for each channel to simulate a mean registration error of n pixels. The decrease in spectral quality was calculated using a set of quality metrics from the resulting misaligned spectral image y, using the original aligned image as the reference x. An example is shown in Fig. 5

Several measures have been proposed for assessing spectral image quality, see, e.g., [43, 44]. The metrics for spectral image y and reference spectral image x adopted for the evaluation were CC, root mean squared error (RMSE), spectral correlation measure (SCM) and spectral information divergence (SID), defined as

$$\mathrm{CC} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \tag{2}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - y_i)^2}{n}}, \tag{3}$$

$$\mathrm{SCM} = \frac{\sum_{i=1}^{d} x_i y_i - \sum_{i=1}^{d} x_i \sum_{i=1}^{d} y_i}{\sqrt{\left[d \sum_{i=1}^{d} x_i^2 - \sum_{i=1}^{d} (x_i)^2 \right] \left[d \sum_{i=1}^{d} y_i^2 - \sum_{i=1}^{d} (y_i)^2 \right]}}, \tag{4}$$

$$\tag{5}$$

$$\begin{aligned} \mathrm{SID} = & \sum_{i=1}^{d} p_i (\log p_i - \log q_i) + \\ & \sum_{i=1}^{d} q_i (\log q_i - \log p_i), \end{aligned} \tag{6}$$

$$p_i = \frac{x_i}{\sum_{j=1}^{d} x_j}, \quad q_i = \frac{y_i}{\sum_{j=1}^{d} y_j},$$

where σ is the standard deviation, n is the number of channel image pixels, d is the number of channels in the spectra.

The parameters for each registration method were systematically selected by registering a subset of the test images whilst varying the parameter values. The parameter combination that produced the smallest error on the images was chosen. All registration errors were measured as the Euclidean distance between the grid points of the registered image and the base image. In the semisynthetic test set, it is possible that the registration method, in addition to estimating the synthetic transform, corrects some of the system error. As this would manifest itself as displacement with respect to the base image location, it would show as increased error despite the more accurate registration. Therefore, error values below the system error of 2.3 pixels ([8]) are considered as zero for the semisynthetic test set.

5.2. Results and discussion

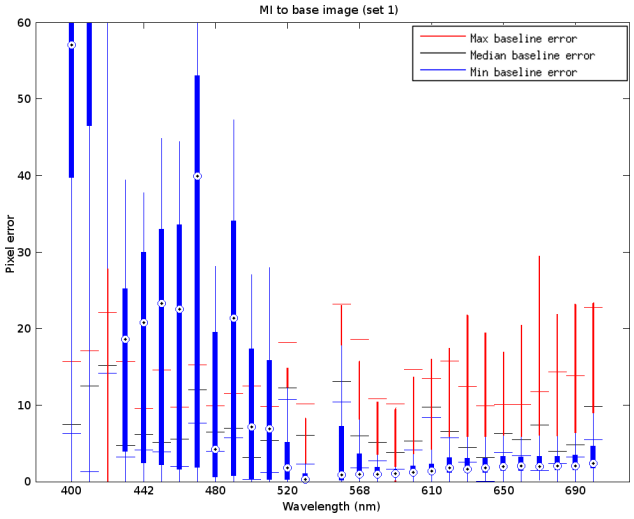
The registration accuracy was measured as pixel error between the grid points resulting from the aligning transformation and the original ones before the known transformation. The amount of distortion applied to the test image sets increased from Set 1 to Set 5, with the distortion in Set 5 corresponding to a 5° offset from the camera axis. The amount of translation applied to the datasets was increased accordingly. The baseline error visualised in the figures refers to the error before registration, i.e., the amount of displacement in a given test image.

MI produced relatively good results for the medium and long wavelengths of the first synthetic set. However, it performed poorly for wavelengths shorter than 520 nm and showed sensitivity to the increasing level of deformation with the successive sets (see Fig. 6). The larger translation in datasets 4 and 5 proved to have a significant impact on the performance. Furthermore, due to the deformable transformation, regions of high registration error could be found in otherwise well registered images.

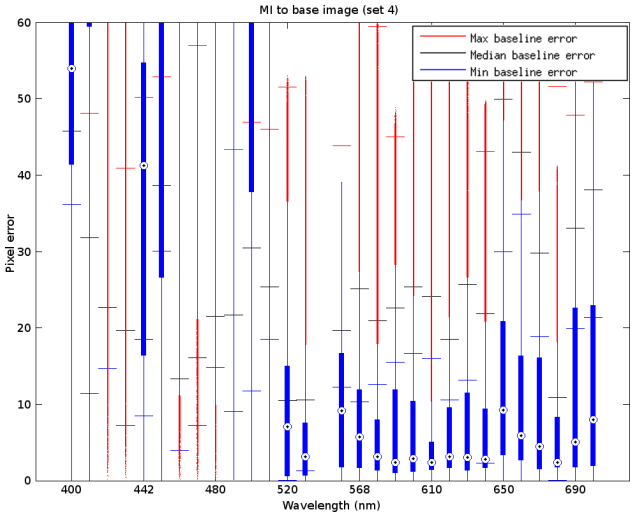
Using intermediate templates with MI significantly decreased the registration error in the shorter wavelengths (see Fig. 7). While the shortest path strategy similarly improved the registrations in the shorter wavelengths, there was an adverse effect to the general registration performance (shown in Fig. 9).

In the semisynthetic set, the intermediate template strategy improved the performance of MI to a point where the mean registration error was within the system error for the majority of the images. The errors are visualised in Fig. 8. However, most of the registered images in the test set contained regions with higher registration error, i.e., the registrations were only partially successful. For the semisynthetic set, MI did not show similar loss of performance with shorter wavelengths.

RC performed similarly to MI for the medium and longer wavelengths, but showed significantly better performance in the shorter wavelengths. However, similar regions of larger error were present (although less severely



(a) Set 1

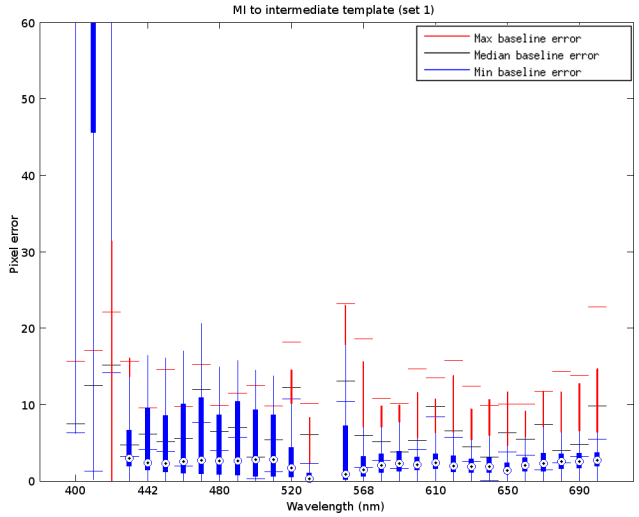


(b) Set 4

Figure 6: mutual information (MI) errors with the synthetic set and all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.

than with MI) in the images, and the method showed similar sensitivity to increasing deformation of the test images. Furthermore, for a few images of the synthetic set, RC failed completely (i.e., the registration error for all pixels was over 60). Neither of the joint registration strategies provided any significant increase in performance for RC. The errors are visualised in Fig. 10.

GDB-ICP showed very good performance for both the synthetic and semisynthetic test sets. With the exception of failed registrations for two pairs of the synthetic set, and one failed registration in Set 5 in the semisynthetic set, the method achieved a reasonable registration error with minimal standard deviation. The median registration error remained below 2 pixels for the majority of



(a) Set 1

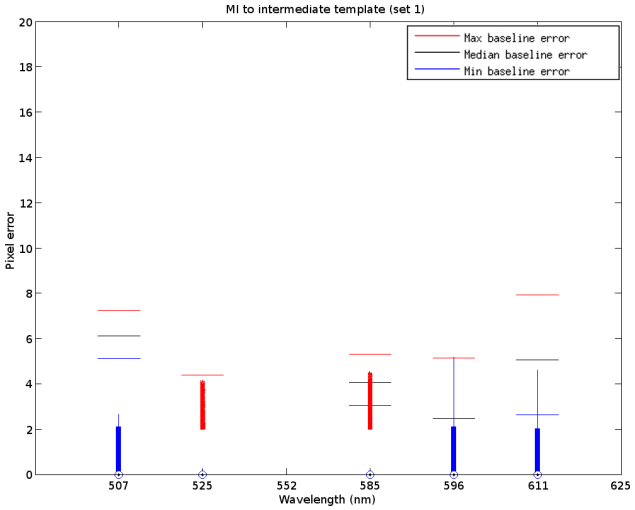
Figure 7: mutual information (MI) errors with the synthetic set and with the intermediate template strategy; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.

the images. As the level of deformation increased in the synthetic images, some regions showed increasing registration error (denoted as outliers). The results are shown in Figures 11 and 12

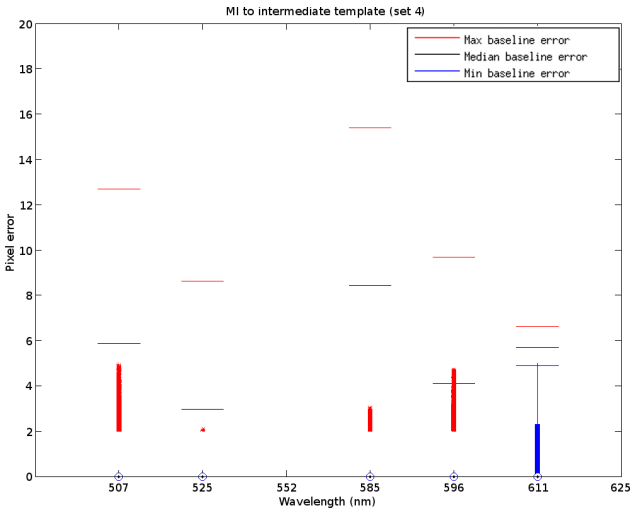
For GDB-ICP, registering to a single base image provided the lowest error for the method in both synthetic and semisynthetic sets. Likely due to restricting the transformation complexity to quadratic, combining of transformations or using intermediate images resulted in a notable decrease in performance.

The intermediate template strategy worked poorly for GDB-ICP. As any displacement due to inaccurate registration accumulates in the template and the quadratic transformation used by GDB-ICP could not fully compensate for the increasing complexity of the accumulated error, the template became blurry and caused further error in consequent registrations.

The strategy of sequential registration did not work well with any of the evaluated methods. The strategy causes additional problems for non-deformable registration approaches (i.e., approaches that limit the complexity of the transformation used for registration). Even if the transformation error is negligible, as a non-rigid transformation is required to properly represent the deformation due to eye movement, combining the sequential registrations results in different channels being registered with transformations of different complexities. In addition, complex transformations require multiple transformation steps as the combined transformation cannot be expressed as a multiplication of the transformation matrices. Quantitatively, the accumulation of registration error significantly overweighted any benefit gained by having less difference



(a) Set 1



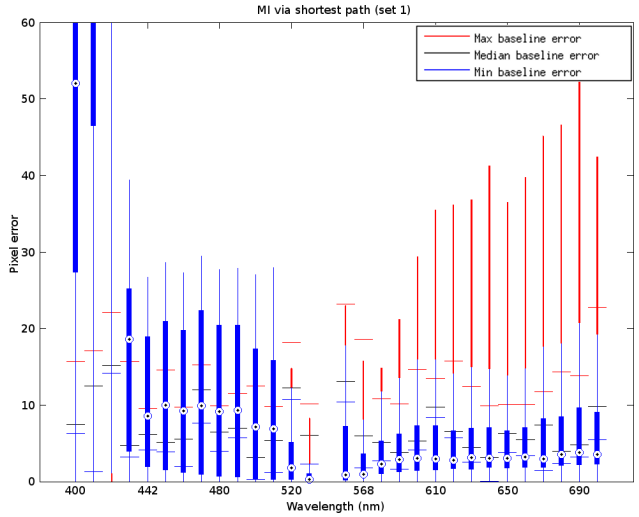
(b) Set 4

Figure 8: mutual information (MI) errors with the semisynthetic set and with the intermediate template strategy; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.

in the image wavelengths.

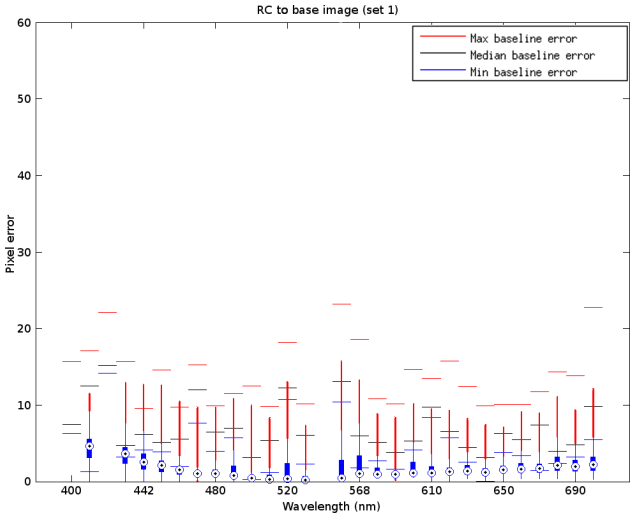
Using the shortest path can help to avoid large wavelength dependent differences between images and limit the accumulation of error due to combined intermediate transforms. However, unless the length of the shortest path is the same for all images, the images will be registered with transformations of different complexity when using non-deformable approaches. The strategy is also sensitive to the choice of the similarity measure used in constructing the cost matrix (i.e., how the distance between images is measured when determining the shortest path). Based on the results, the shortest path strategy is not an effective strategy for the joint registration of spectral channel images.

In general, with the exception of using intermediate

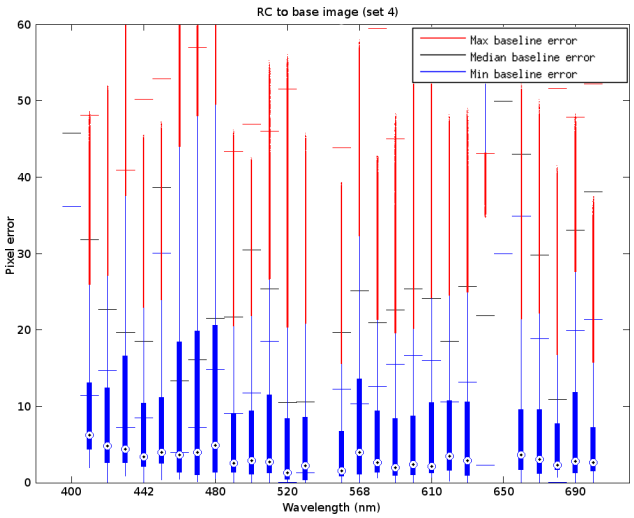


(a) Set 1

Figure 9: mutual information (MI) errors with the synthetic set and with the shortest path strategy; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.



(a) Set 1



(b) Set 4

Figure 10: minimization of residual complexity (RC) errors with the synthetic set and all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.

templates with MI, the evaluated registration strategies showed limited benefit. The increase in registration error due to even limited number of combined transformations outweighed the benefit of more similar image content due to smaller difference in wavelength.

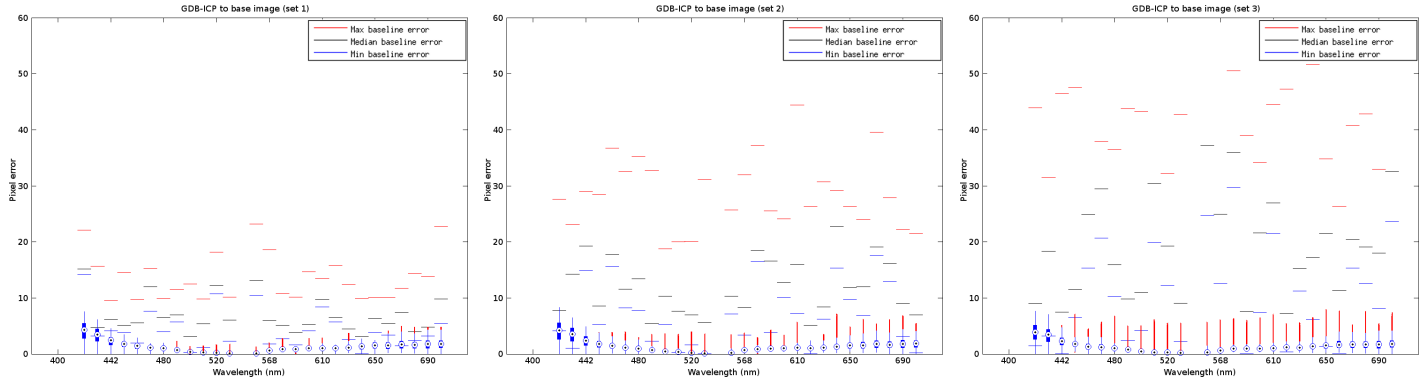
The results of the image registration performance comparison on the synthetic data are presented in Table 2 and summarized in Fig. 13. Similarly, the results on the semisynthetic data are presented in Table 3 and summarized in Fig. 14.

The registration errors tended to be more concentrated towards the image edges opposite to the optic disk (surrogate) where the blood vessels become less prominent. The retinal background was largely unable to provide either features for GDB-ICP or reliable region matches for MI or RC. The lowest registration errors were generally found in regions with high-contrast retinal blood vessels.

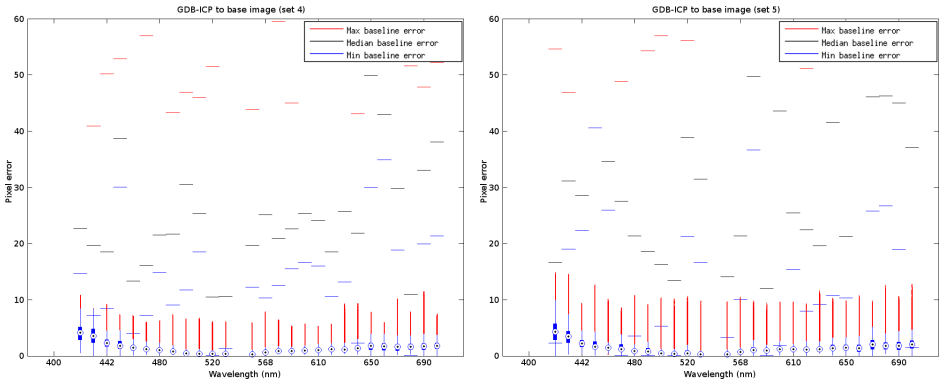
The low error rate of GDB-ICP was partially due to the restriction to the transformation complexity. The global transformation ensured that the registration error was reasonably low even in regions where there were few retinal features to guide the registration process. In contrast, MI and RC generated, for some images, transformations where the parts of an image containing well defined features were registered to sub-pixel accuracy and the registration errors of tens of pixels could be found elsewhere. The effect is visualised in Fig. 15 and Fig. 16.

After a sharp decrease in spectral quality between the original aligned image and image with displacement of two pixels on each channel (the interpolation of image values is likely a significant factor in the decrease of spectral quality), the spectral quality shows close to linear decrease with the increased displacement. While the registration errors remaining in a spectral image composed with GDB-ICP are unlikely to consist only of translation, its spectral quality value seems to correspond to an image with the same mean displacement in pure translation. The effect of registration error on the quality of image spectra is visualised in Fig. 17.

Unless a threshold below which the registration error cannot be distinguished from the error inherent in the imaging system, such as in the case of the semisynthetic test set, can be determined, a clear distinction between a successful and a failed registration is application dependent and not trivial. In the case of spectral images, determining an error threshold for a successful registration can



(a) Set 1 (b) Set 2 (c) Set 3



(d) Set 4 (e) Set 5

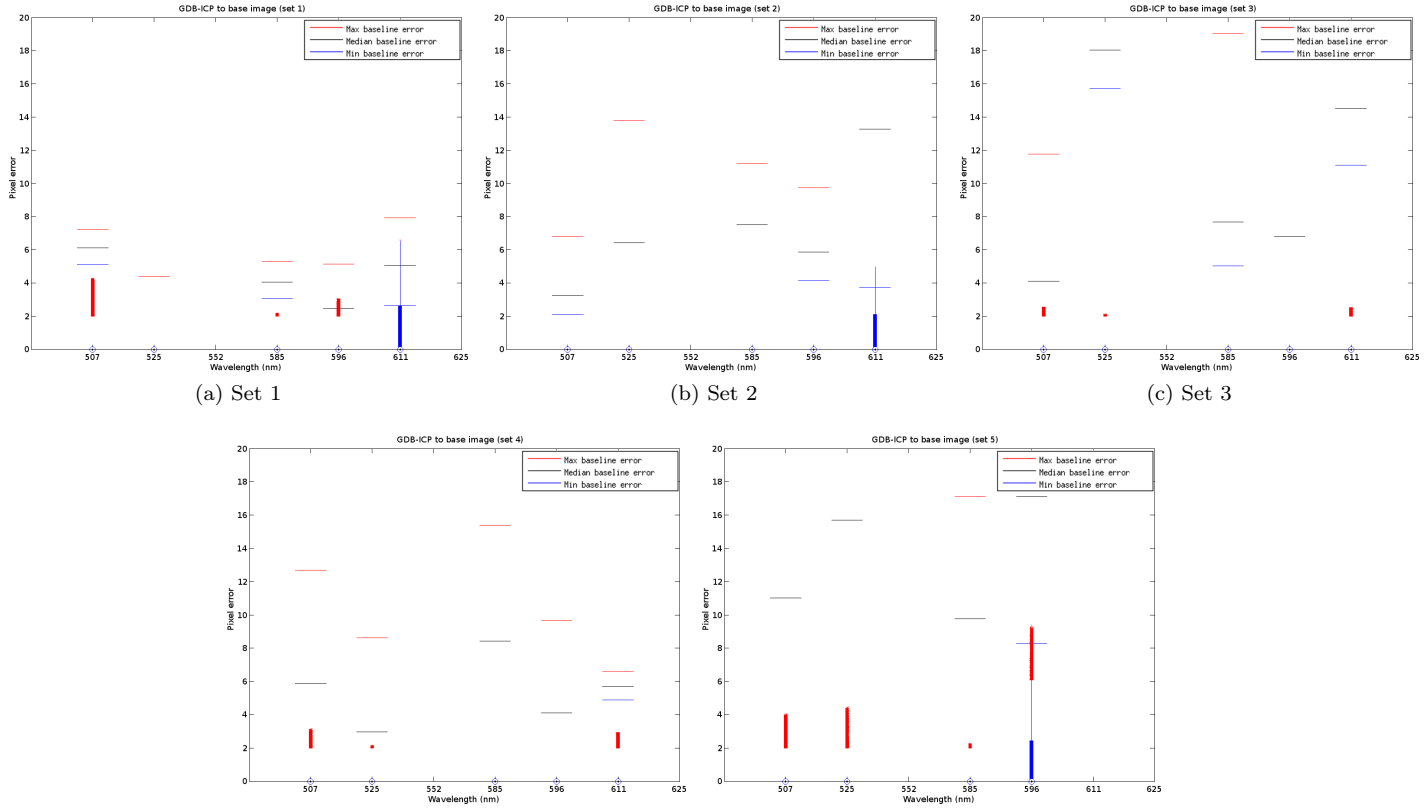
Figure 11: generalized dual-bootstrap iterative closest point (GDB-ICP) errors with the synthetic set and all images registered to a single base image captured at 540 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.

Table 2: Median (Med) and standard deviation (STD) of registration error for the synthetic sets. Init. stands for the initial error before registration, NN for sequential registration, T for intermediate template and SP for shortest path. Lowest error for each method/set is displayed in bold.

Method		Init.	GDB-ICP				MI				RC			
Strategy		All	Base	NN	T	SP	Base	NN	T	SP	Base	NN	T	SP
Med	Set 1	6.1	1.1	1.2	5.5	**1.0**	**2.1**	9.1	2.4	3.5	1.3	**1.2**	1.6	1.4
	Set 2	11.9	**1.1**	2.0	8.5	1.7	**2.3**	18.9	6.0	4.4	**1.5**	3.0	1.7	1.9
	Set 3	18.0	**1.2**	5.2	11.5	1.4	20.4	186.2	25.0	**4.3**	**2.2**	10.5	3.0	2.3
	Set 4	22.7	**1.2**	4.6	22.3	2.4	8.0	141.1	18.0	**7.1**	**3.0**	7.7	3.9	3.9
	Set 5	31.1	**1.2**	8.2	47.3	2.2	47.5	275.9	**42.7**	58.3	**4.6**	35.1	7.1	5.3
STD	Set 1	1.6	**0.4**	0.7	1.8	0.5	4.1	15.1	**2.5**	8.9	1.5	2.4	**1.3**	1.4
	Set 2	3.4	**0.5**	2.1	3.3	1.0	5.8	33.8	**5.1**	10.9	3.6	4.7	3.1	**3.0**
	Set 3	6.2	**0.6**	3.7	5.1	0.9	19.3	73.2	17.8	**15.8**	**5.8**	9.9	6.2	**5.8**
	Set 4	7.6	**0.8**	4.0	8.0	1.7	13.3	54.4	**8.9**	19.3	8.4	11.5	**7.4**	8.5
	Set 5	10.9	**0.9**	8.2	6.3	1.8	**29.4**	83.2	29.7	30.8	12.6	13.4	10.8	**10.2**

be approached through the effect of the registration inaccuracy on the spectra; a registration error of two pixels would require the spatial resolution of the spectral image to be downscaled by a factor of two to provide accurate spectra at each pixel.

The system by Fält et al. [4] is capable of acquiring images at 2048×2048 resolution. Therefore, a registration error of two or three pixels would still produce images of high spatial resolution and with accurate spectrum at each pixel. GDB-ICP is capable of registering most images with the median error within this accuracy. For the first set, RC achieves similar median error, but some images contain regions with considerably higher errors. The large errors of MI in the blue wavelength images make the method unreliable in spectral image composition.



(a) Set 1 (b) Set 2 (c) Set 3

(d) Set 4 (e) Set 5

Figure 12: generalized dual-bootstrap iterative closest point (GDB-ICP) errors with the semisynthetic set and all images registered to a single base image captured at 552 nm; the median error is shown with a circle, the boxes represent the 25th and 75th percentiles, and the whiskers extend to the most extreme values still considered as inliers. The outliers are plotted individually. Baseline error is the error before registration.

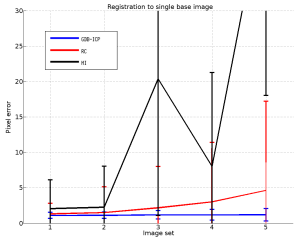
Table 3: Median (Med) and standard deviation (STD) of registration error for the semisynthetic sets. Baseline stands for the initial error before registration, NN for sequential registration, T for intermediate template and SP for shortest path. Lowest error for each method/set is displayed in bold.

Method		Init.	GDB-ICP				MI				RC			
Strategy		All	Base	NN	T	SP	Base	NN	T	SP	Base	NN	T	SP
Med	Set 1	4.1	0.0	0.0	0.0	0.0	4.8	4.6	**0.0**	4.6	**2.6**	4.0	3.3	**2.6**
	Set 2	6.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Set 3	7.7	0.0	0.0	0.0	0.0	3.0	4.2	**0.0**	3.5	**0.0**	**0.0**	11.0	**0.0**
	Set 4	5.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Set 5	15.7	**0.0**	2.4	7.3	2.3	**2.1**	3.2	10.2	3.1	**2.5**	3.6	5.4	2.9
STD	Set 1	1.0	0.9	0.6	**0.5**	0.7	2.9	2.4	**1.0**	2.8	**3.4**	5.2	4.3	3.8
	Set 2	2.7	**0.0**	**0.0**	1.4	**0.0**	**1.2**	1.5	1.5	1.4	1.3	1.3	3.0	**0.9**
	Set 3	3.0	**0.1**	0.8	0.8	0.5	**1.9**	2.5	2.8	2.1	**1.4**	7.1	6.7	3.0
	Set 4	2.1	0.1	**0.0**	1.3	**0.0**	1.0	0.8	0.9	**0.5**	1.0	**0.8**	1.3	**0.8**
	Set 5	5.0	**0.9**	1.8	5.0	1.5	**2.1**	4.7	8.6	4.1	4.4	**3.8**	7.7	**3.8**

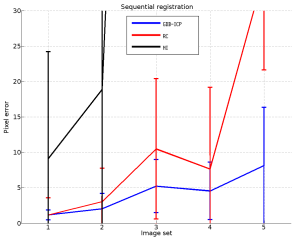
6. Conclusion

This paper focused on the comparison of image registration methods generalized dual-bootstrap iterative closest point (GDB-ICP), mutual information (MI) and minimization of residual complexity (RC) which were selected based on a previous study [11]. In addition, image set registration strategies were studied for composing spectral retinal images. The experiments on five sets of channel images of an eye phantom and artificially deformed medical data (synthetic and semisynthetic data) showed that the registration error increases with increasing wavelength difference between the floating and base image.

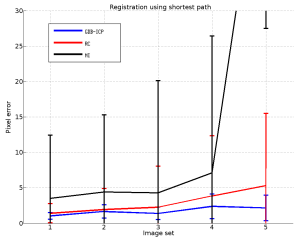
GDB-ICP outperformed the other methods in registration accuracy and achieved low registration error for most



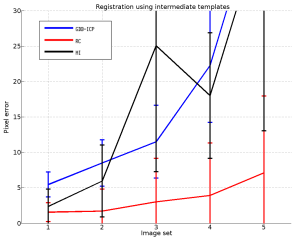
(a) Registration to single base image.



(b) Sequential registration.

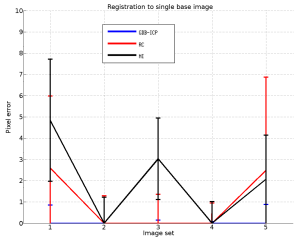


(c) Shortest path registration.

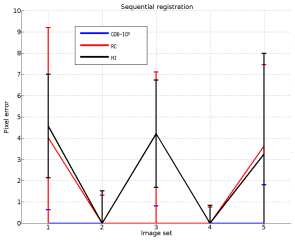


(d) Template registration.

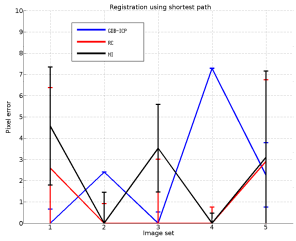
Figure 13: Registration errors with the synthetic sets and different image set registration strategies; the single base image was captured at 540 nm; the median error is shown with the standard deviation as the vertical bar.



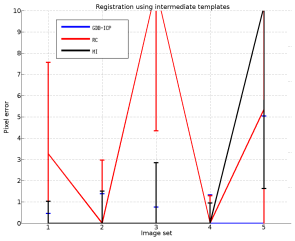
(a) Registration to single base image.



(b) Sequential registration.



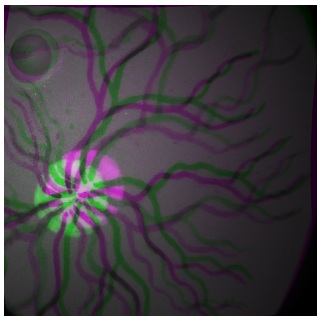
(c) Shortest path registration.



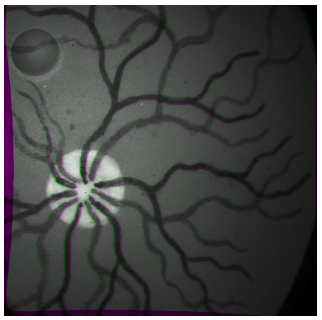
(d) Template registration.

Figure 14: Registration errors with the semisynthetic sets and different image set registration strategies; the single base image was captured at 552 nm; the median error is shown with the standard deviation as the vertical bar.

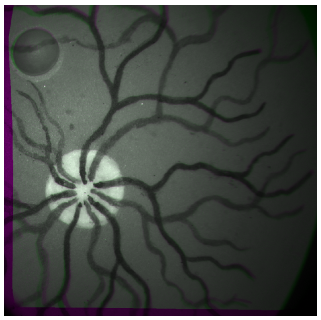
of the synthetic and semisynthetic test sets. Different registration strategies did not provide consistent improvement to the performance of GDB-ICP compared to registering all channel images to a single base image. Based on the experiments, the increase in mean registration error has a close-to-linear relationship to the quality of spectra in the composed spectral image.



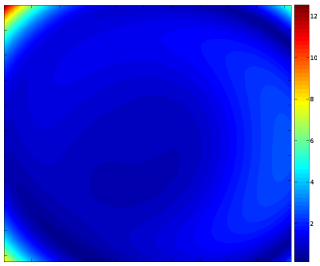
(a) Difference image (unaligned).



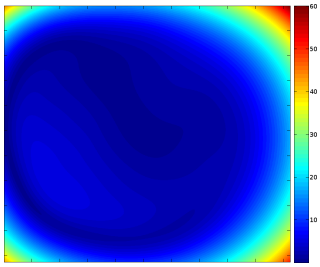
(b) Difference image (GDB-ICP).



(c) Difference image (RC).



(d) Error map (GDB-ICP).



(e) Error map (RC).

Figure 15: Examples of the spatial distribution of error; (a)-(c) show false colour difference images of registration result and ground truth.

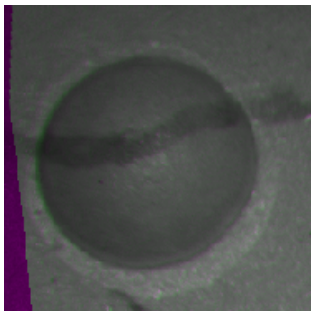
Acknowledgements

The authors would like to thank the Academy of Finland for the financial support of the ReVision project (No. 259560).

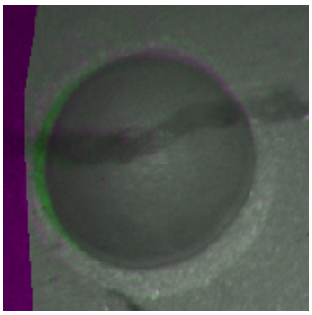
References

[1] E. Stefánsson, T. Bek, M. Porta, N. Larsen, J. K. Kristinsson, E. Agardh, Screening and prevention of diabetic blindness, Acta Ophthalmologica Scandinavica 78 (4) (2000) 374–385.

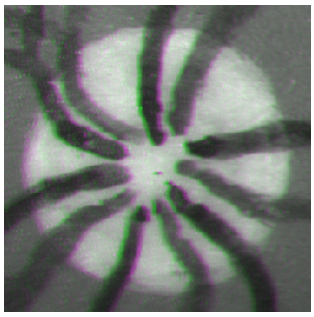
[2] G. Scotland, P. McNamee, A. Fleming, K. Goatman, S. Philip, G. Prescott, P. Sharp, G. Williams, W. Wykes, G. Leese, J. Olson, Costs and consequences of automated algorithms versus manual grading for the detection of referable diabetic retinopathy, British Journal of Ophthalmology 94 (6) (2010) 712–719.



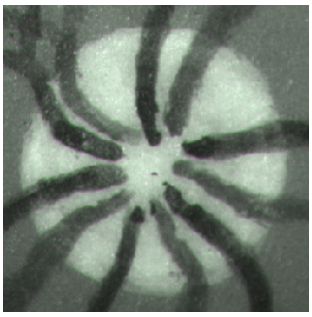
(a) Fovea region difference image (GDB-ICP).



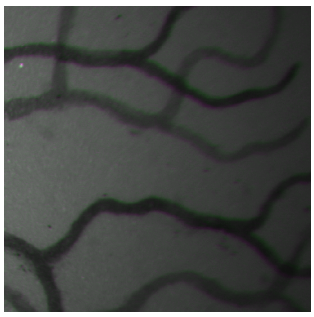
(b) Fovea region difference image (RC).



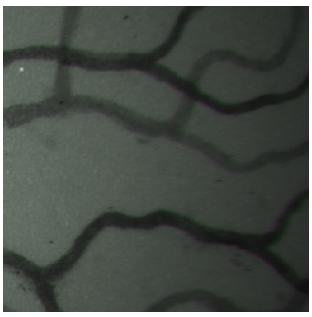
(c) Optic disc region difference image (GDB-ICP).



(d) Optic disc region difference image (RC).



(e) Vessel region difference image (GDB-ICP).



(f) Vessel region difference image (RC).

Figure 16: Examples of regions with higher spatial error; (a)-(f) show false colour difference images of registration result and ground truth.

[3] T. T. J. M. Berendschot, P. J. DeLint, D. van Norren, Fundus reflectance – historical and present ideas, Progress in Retinal and Eye Research 22.

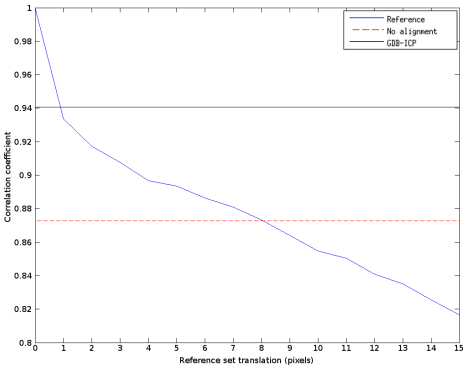
[4] P. Fält, J. Hiltunen, M. Hauta-Kasari, I. Sorri, V. Kalesnykiene, J. Pietilä, H. Uusitalo, Spectral imaging of the human retina and computationally determined optimal illuminants for diabetic retinopathy lesion detection, Journal of Imaging Science and Technology 55 (3) (2011) 030509–1–030509–10.

[5] L. Gao, R. T. Smith, T. S. Tkaczyk, Snapshot hyperspectral retinal camera with the image mapping spectrometer (ims), Biomedical Optics Express 3 (1) (2012) 48–54.

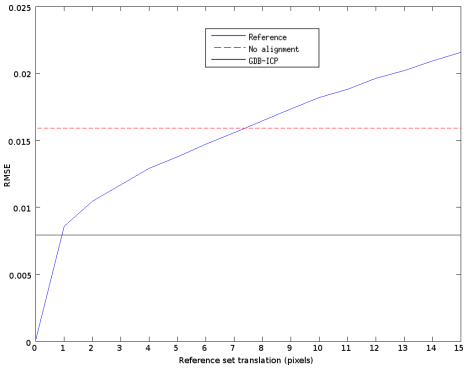
[6] W. R. Johnson, D. W. Wilson, W. Fink, M. Humayun, G. Bearman, Snapshot hyperspectral imaging in ophthalmology, Journal of Biomedical Optics 12 (1) (2007) 014036.

[7] B. Khoobehi, J. M. Beach, H. Kawano, Hyperspectral imaging for measurement of oxygen saturation in the optic nerve head, Investigative Ophthalmology & Visual Science 45 (2004) 1464–1472.

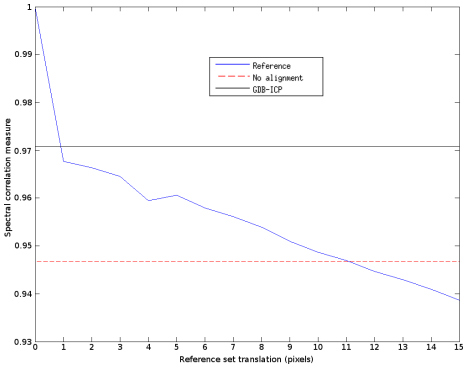
[8] I. Styles, A. Calcagni, E. Claridge, F. Orihuela-Espina, J. Gib-



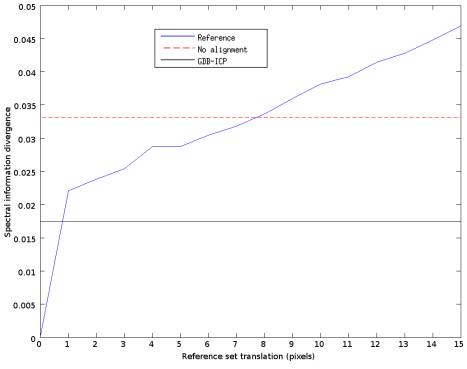
(a) Correlation coefficient.



(b) Root mean squared error.



(c) Spectral correlation measure.



(d) Spectral information divergence.

Figure 17: The effect of registration error on the quality of image spectra.

son, Quantitative analysis of multi-spectral fundus images, Medical Image Analysis 10 (4) (2006) 578–597.
[9] P. Viola, W. Wells III, Alignment by maximization of mutual information, International Journal of Computer Vision 24 (2) (1997) 137–154.
[10] J. Zheng, J. Tian, K. Deng, X. Dai, X. Zhang, M. Xu, Salient feature region: a new method for retinal image registration, IEEE Transactions on Information Technology in Biomedicine 15 (2) (2011) 221 – 232.
[11] L. Laaksonen, E. Claridge, P. Fält, M. Hauta-Kasari, H. Uusitalo, L. Lensu, Comparison of image registration methods for composing spectral retinal images, in: Proceedings of International Workshop on Ophthalmic Medical Image Analysis, Boston, USA, 2014, pp. 57–64.
[12] B. Zitova, J. Flusser, Image registration methods: A survey, Image and Vision Computing 21 (11) (2003) 977–1000.
[13] M. Wyawahare, P. Patil, H. Abhyankar, Image registration techniques: An overview, International Journal of Signal Processing, Image Processing and Pattern Recognition 2 (3) (2009) 11–28.
[14] J. Maintz, M. Viergever, A survey of medical image registration, Medical Image Analysis 2 (1) (1998) 1–36.
[15] D. Hill, P. Batchelor, M. Holden, D. Hawkes, Medical image registration, Physics in Medicine and Biology 46 (3) (2001) R1.
[16] K. Bhatia, J. Hajnal, A. Hammers, D. Rueckert, Similarity metrics for groupwise non-rigid registration, in: Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2007, pp. 544–552.
[17] P. Markelj, D. Tomaževič, B. Likar, F. Pernuš, A review of 3d/2d registration methods for image-guided interventions, Medical Image Analysis 16 (3) (2012) 642–661.
[18] F. P. Oliveira, J. M. R. Tavares, Medical image registration: a review, Computer Methods in Biomechanics and Biomedical Engineering 17 (2) (2014) 73–93. doi:10.1080/10255842.2012.670855.
[19] W. Crum, T. Hartkens, D. Hill, Non-rigid image registration: theory and practice, The British Journal of Radiology 77 (2004) S140–S153.
[20] A. Sotiras, C. Davatzikos, N. Paragios, Deformable medical image registration: A survey, IEEE Transactions on Medical Imaging 32 (7) (2013) 1153–1190.
[21] H. Stone, R. Wolpov, Blind cross-spectral image registration using prefiltering and fourier-based translation detection, IEEE Transactions on Geoscience and Remote Sensing 40 (3) (2002) 637–650.
[22] M. Pfingsthorn, A. Birk, S. Schwertfeger, H. Bülow, K. Pathak, Maximum likelihood mapping with spectral image registration, in: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2010, pp. 4282–4287.
[23] M. Hasan, X. Jia, A. Robles-Kelly, J. Zhou, M. R. Pickering, Multi-spectral remote sensing image registration via spatial relationship analysis on sift keypoints, in: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, 2010, pp. 1011–1014.
[24] T. Teng, M. Lefley, D. Claremont, Progress towards automated diabetic ocular screening: a review of image analysis and intelligent systems for diabetic retinopathy, Medical and Biological Engineering and Computing 40 (1) (2002) 2–13.
[25] F. Laliberté, L. Gagnon, Y. Sheng, Registration and fusion of retinal images - an evaluation study, IEEE Transactions on Medical Imaging 22 (5) (2003) 661–673.
[26] GDB-ICP, `http://www.vision.cs.rpi.edu/gdbicp/`, accessed: 11 June 2014.
[27] Diffeomorphic demons, `http://se.mathworks.com/matlabcentral/fileexchange/21451-multimodality-non-rigid-demon-algorithm-image-registration`, accessed: 11 June 2014.
[28] Diffeomorphic log demons, `https://se.mathworks.com/matlabcentral/fileexchange/39194-diffeomorphic-log-demons-image-registration`, accessed: 11 June 2014.
[29] Medical image registration toolbox, `MedicalImageRegistrationToolbox`, accessed: 11 June 2014.
[30] A. Myronenko, X. Song, Intensity-based image registration by minimizing residual complexity, IEEE Transactions on Medical Imaging 29 (11) (2010) 1882–1891.
[31] G. Yang, C. Stewart, M. Sofka, C.-L. Tsai, Registration of challenging image pairs: Initialization, estimation and decision, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (11) (2007) 1973 – 1989.
[32] Z. Xue, K. Wong, S. Wong, Joint registration and segmentation of serial lung ct images for image-guided lung cancer diagnosis and therapy, Computerized Medical Imaging and Graphics 34 (1) (2010) 55–60.
[33] G. Wu, Q. Wang, H. Jia, D. Shen, Feature-based groupwise registration by hierarchical anatomical correspondence detection, Human Brain Mapping 33 (2) (2012) 253–271.
[34] K. Bhatia, P. Aljabar, J. Boardman, L. Srinivasan, M. Murgasova, S. Counsell, M. Rutherford, J. Hajnal, A. Edwards, D. Rueckert, Groupwise combined segmentation and registration for atlas construction, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007, Springer, 2007, pp. 532–540.
[35] C. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, C. Taylor, A unified information-theoretic approach to groupwise non-rigid registration and model building, in: Information Processing in Medical Imaging, Springer, 2005, pp. 1–14.
[36] S. Balci, P. Golland, W. Wells, Non-rigid groupwise registration using b-spline deformation model, Open Source and Open Data for MICCAI (2007) 105–121.
[37] S. Joshi, B. Davis, M. Jomier, G. Gerig, Unbiased diffeomorphic atlas construction for computational anatomy, NeuroImage 23 (2004) S151–S160.
[38] J. Hamm, D. Ye, R. Verma, C. Davatzikos, Gram: A framework for geodesic registration on anatomical manifolds, Medical Image Analysis 14 (5) (2010) 633–642.
[39] H. Jia, G. Wu, Q. Wang, D. Shen, Absorb: Atlas building by self-organized registration and bundling, NeuroImage 51 (3) (2010) 1057–1070.
[40] A. Calcagni, J. M. Gibson, I. B. Styles, E. Claridge, F. Orihuela-Espina, Multispectral retinal image analysis: a novel non-invasive tool for retinal imaging, Eye 25 (12).
[41] I. Bekerman, P. Gottlieb, M. Vaiman, Variations in Eyeball Diameters of the Healthy Adults, Journal of Ophthalmology 2014 (2014) 1–5. doi:10.1155/2014/503645.
URL `http://www.hindawi.com/journals/joph/2014/503645/`
[42] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, Speeded-up robust features (surf), Computer Vision and Image Understanding 110 (3) (2008) 346–359.
[43] A. Vaiopoulos, Developing matlab scripts for image analysis and quality assessment, in: SPIE Remote Sensing, International Society for Optics and Photonics, 2011, pp. 81810B–81810B.
[44] F. van der Meer, The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery, International Journal of Applied Earth Observation and Geoinformation 8 (1) (2006) 3–17.